Heterogeneity assumptions in the specification of bargaining models: a study of household level trade-offs between commuting time and salary

Vikki O'Neill[*] *Stephane Hess* [†]

Abstract

With many real world decisions being made in conjunction with other decision makers, or single agent decisions having an influence on other members of the decision maker's immediate entourage, there is strong interest in studying the relative weight assigned to different agents in such contexts. In the present paper, we focus on the case of one member of a two person household being asked to make choices affecting the travel time and salary of both members. We highlight the presence of significant heterogeneity across individuals not just in their underlying sensitivities, but also in the relative weight they assign to their partner, and show how this weight varies across attributes. This is in contrast to existing work which uses weights assigned to individual agents at the level of the overall utility rather than for individual attributes. We also show clear evidence of a risk of confounding between heterogeneity in marginal sensitivities and heterogeneity in the weights assigned to each member. We show how this can lead to misleading model results, and argue that this may also explain past results showing bargaining or weight parameters outside the usual $[0, 1]$ range in more traditional joint decision making contexts. In terms of substantive results, we find that male respondents place more weight on their partner's travel time, while female respondents place more weight on their partner's salary.

Keywords: household decisions; distributional assumptions; random coefficients; joint decisions; bargaining coefficient

[*]Medical Research Council Biostatistics Unit, Institute of Public Health, Cambridge, CB2 0SR, vikki.oneill@mrc-bsu.cam.ac.uk

[†]Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, s.hess@its.leeds.ac.uk

1 Introduction

Data on choice behaviour is routinely used to derive individuals' preferences for goods and services. However, there is acknowledgment across fields that many real life decisions are made not by a single person in isolation, but in consultation with other actors. Similarly, a single person may make choices that affect other members of their household or peer group. The majority of such work has looked at decisions in a household context, and this will be the framework for the remainder of this paper.

If choices are made jointly by a number of household members, then it is likely that they take part in a negotiation process in order to maximise some joint-utility function. Similarly, when an individual is making a decision that will affect more than just themselves, the expectation is that, at least to some degree, they will take into consideration the preferences held by other household members (or perceived to be held), which may be different from theirs. They are also likely to give differential weight to their own preferences and those of their partner across different attributes.

In the context of joint decisions, the recognition of the differential influence of individual players has moved us away from the unitary household model or 'common preference model' which assumes that, irrespective of the members of a household, it will act as a single-decision-making unit, wherein a single preference function will represent all members of the group (see, for example, discussions in Adamowicz et al., 2005, Katz, 1997, Lampietti, 1999 and Vermeulen, 2002). This has led to a significant body of work looking at how members of a household may engage in a process of joint deliberation in order to maximise both their individual and joint utility functions (see, for example Adamowicz et al., 2005, Marcucci et al., 2011 and Munro, 2009 for a comprehensive review, as well as key developments in Aribarg et al., 2002, Arora and Allenby, 1999, Browning and Chiappori, 1998, Dellaert et al., 1998, Dosman and Adamowicz, 2006 and Hensher et al., 2008). Within this literature, it is evident that there is not only disparity between household member's preferences, but also between the choices made by individuals and the choices made by households collectively.

While some analysts have explicitly modelled the bargaining process (Hensher et al., 2008), this requires a very specific approach to data collection, using an iterative process. In the majority of work however, only information on choices is observed, as the bargaining process is not *captured* explicitly in data collection (Dosman and Adamowicz, 2006). The key here is that choices are observed for individual respondents in addition to the joint choices, and that estimation on a pooled dataset allows the calibration of weights attached to individual decision makers, which represent their influence in the joint choice. An important area of interest in that context has been the study of heterogeneity across respondents, both in terms of their sensitivities, as well as their *weight* in this bargaining process (see e.g. Beharry-Borg et al., 2009). Crucially, this model approach is suitable not just for the analysis of joint decisions, but also the analysis of data where one respondent makes choices affecting multiple agents. The work described in the present paper falls into this last category.

In common with work for example by the above cited Beharry-Borg et al. (2009), the present paper makes the case that, just as in more traditional choice data (i.e. choices by a single agent affecting only themselves), there exist significant differences across people in the context of household level decisions. Our assertion is that not adequately representing such heterogeneity, both in the underlying sensitivities and the relative weight assigned to a person's own sensitivities and those of their partner, may lead to misguided findings. Crucially, there is significant risk of confounding between heterogeneity in the marginal utility coefficients and the bargaining or weight parameters, where inappropriate specifications are likely to exacerbate problems. We also argue that there may be heterogeneity across attributes in the weights assigned to individual agents, thus highlighting the potential disadvantages of the common assumption in the literature that the relative importance of an agent is constant across attributes.

We support these claims through an empirical analysis using stated choice data examining the intra-household preferences for commuting time and salary collected in the Stockholm region of Sweden. Specifically, in this survey, each member of a dyadic[1] household was individually asked to trade between their own commuting time and salary and also their partner's commuting time and salary. While the emphasis in this paper is on decisions at the household level, the methodological discussions clearly also have relevance in other joint decision-making contexts relying on the *bargaining* model. Similarly, even though in contrast to the recreational choice contexts of Dosman and Adamowicz (2006) and Beharry-Borg et al. (2009), our work looks at the choice to travel to work, the modelling framework is general and applies across contexts.

Our results suggest the presence of significant levels of heterogeneity both in the underlying

[1] A household containing two individuals, living as partners.

sensitivities of individual respondents as well as in the weights they assign to their partners. A failure to jointly account for both types of heterogeneity leads to inferior results and possibly misguided interpretations. Furthermore, either not accounting jointly for the heterogeneity in the utility and weight parameters, or making inappropriate distributional assumptions, or using utility rather than attribute level weight parameters, can play a strong role in producing results that indicate weight parameters outside the $[0, 1]$ range. We argue that our theoretical claims and empirical results in part explain such results in previous work.

The specific contribution of this paper is thus to highlight the interaction between the heterogeneity assumptions for the utility parameters and bargaining or weight coefficients, and to make the case for attribute specific rather than utility level weights for the individual decision makers. Although existing work has looked at the issue of taste heterogeneity and has allowed either for deterministic (Dosman and Adamowicz, 2006) or random heterogeneity (Beharry-Borg et al., 2009) in the weight parameters, it has not adequately addressed the issues of confounding and the impact of distributional assumptions. Additionally, while attribute specific weight parameters are referred to by Beharry-Borg et al. (2009), their estimation still relies on utility level weight parameters, further increasing the novelty of our work.

The remainder of this paper is organised as follows. Section 2 presents an overview of the models that are applicable in this context, with a particular emphasis on the specification of bargaining or weight parameters. This is followed by our empirical application in Section 3, and a concluding discussion is presented in Section 4.

2 Theory

Independently of whether the choice relates to a joint decision or a single person making a decision for a household, the utility that household h obtains from choosing alternative j is represented as:

$$U_{hj} = V_{hj} + \varepsilon_{hj}, \tag{1}$$

where V_{hj} is the deterministic component of utility and ε_{hj} is the random component. Focussing on a two-person context, we recognise that the different members of a household potentially have different marginal sensitivities (i.e. we have β_1 for person 1 and β_2 for person 2), carry different weight in a joint decision process or are given different weight by the person making decisions affecting both people. As such, we now have that:

$$V_{hj} = \lambda_1 f(\beta_1, x_{1j}) + \lambda_2 f(\beta_2, x_{2j}), \tag{2}$$

where x_{1j} and x_{2j} relate to the vector x of explanatory variables for alternative j for the two household members. The functional form of the utility function is defined by $f(\beta_1, x_{1j})$, where the majority of applications rely on a linear in parameters specification. The two additional parameters λ_1 and λ_2 give the weights of the two household members (either in the joint decision making process or differences in the weight assigned by the single decision maker), where we have that $\lambda_1 + \lambda_2 = 1$ for identification reasons. Usually, the assumption is also made that $0 \leq \lambda_p \leq 1$, $p = 1,2$, a point we will return to below.

Existing work has relied on generic λ parameters across attributes, thus assuming that the weight assigned to a given agent is constant across attributes. This is clearly a simplistic assumption which is derived in particular from the notion of influence of one person in a joint decision making process but which does not recognise that the influence of given agents may vary across attributes. This possibility was acknowledged by Beharry-Borg et al. (2009) but not used in their estimations. Again without making assumptions about functional form, Equation 2 would be replaced by:

$$V_{hj} = \sum_{k=1}^{K} \lambda_{1,k} f_k(\beta_{1,k}, x_{1j,k}) + \lambda_{2,k} f_k(\beta_{2,k}, x_{2j,k}), \tag{3}$$

where the subscript k now refers to attribute k out of K.

A model of the type shown in Equation 2 or Equation 3 needs to be estimated on pooled data containing individual choices as well as either joint choices or choices affecting both agents but made by one respondent. The joint estimation of both β_1, β_2 and λ_p is only possible when individual choices are observed for both agents, in addition to joint choices. When the choices affecting both agents are made by one respondent only, who also provides individual choices affecting only the respondent himself or herself, then we can either estimate β_1 and β_2, or β and λ_p. With the relevance of the model specification to data on joint choices in mind, we make use of the latter in our application[2].

In a model estimated on data with joint choices, λ seeks to capture the influence that each decision maker has on forming the joint utility function, either overall or at the attribute level. In a model estimated on data containing household choices made by one decision maker, λ is likely to capture both the relative importance that this person attaches to the members of the household, as well as this respondent's perception of the value that their partner would place on the attribute, relative to the decision maker's perception, in the case where attribute specific λ parameters are used.

A significant amount of research has gone into the specification of the λ parameters in such models. The assumption of $\lambda_1 = \lambda_2 = 0.5$ is generally rejected on theoretical as well as empirical grounds. With the weights being freely estimated rather than constrained to be equal, an important question then arises as to the range for these weights. Although it seems reasonable to think that joint taste intensities or household level sensitivities *selected* by one person, should be intermediate between individual taste intensities, i.e. λ falling within the $[0, 1]$ range, this may not always be the case (cf. Adamowicz et al., 2005), and there are examples of estimates outside this range (see, for example Beharry-Borg et al., 2009).

A number of interpretations for a λ estimate outside the $[0, 1]$ interval have been put forward. For instance, Dellaert et al. (1998) describes a negative value for λ as the "systematic denial of the individual's preference in the joint evaluation", whilst Beharry-Borg et al. (2009) suggest that when an individual is a member of a group, their preferences may be even stronger than their individual responses would have been if they were not part of the group. This is known as the group polarization phenomenon (cf. Arora and Allenby, 1999; Myers and Lamm, 1976; Rao and Steckel, 1991; Steckel et al., 1991). Similarly, Bateman and Munro (2005) find couples making more risk adverse choices when facing tasks together compared to when the partners faced the same decision-making tasks individually.

A key hypothesis put forward in the present paper is that λ parameters outside the $[0, 1]$

[2] It can be seen that a model with attribute specific λ_p parameters is equivalent to a model estimating β_1 and β_2, a point we will return to later in the paper.

interval (cf. Dosman and Adamowicz, 2006; Beharry-Borg et al., 2009) may be caused in part by inappropriate specifications and confounding. In particular, we argue that there is scope for heterogeneity in both the utility parameters β and the weight parameters λ, be it deterministic or random heterogeneity, in line with Dosman and Adamowicz (2006); Beharry-Borg et al. (2009). Additionally, we put forward the notion that the weight of individual decision makers varies across attributes, where this could be accommodated in attribute specific λ parameters. Not accounting fully for the heterogeneity across respondents in β and λ as well as the hetereogeneity across attributes in λ not only risks leading to inferior model performance but might cause confounding that could explain some of the previous findings of λ parameters outside the $[0, 1]$ interval. The same clearly applies to using inappropriate distributions for λ which would impose a non-zero probability of values outside the $[0, 1]$ interval rather than allowing them to be retrieved in the analysis. For that reason, we make the case that the bounds on λ should be estimated, rather than imposed, including through using unbounded distributions.

3 Empirical application: a work place location study in Sweden

This section presents the results from our case study of the role of heterogeneity in sensitivities and weights assigned to household members in the scenario where both members of a dyadic household individually provide choices in settings that would affect both members. We first discuss the data before turning our attention to model results, where we initially focus on model specification and results for structures without heterogeneity across respondents before turning to model specification and results for structures allowing for such heterogeneity.

3.1 Data

The data used for this application come from a survey conducted in the Stockholm region of Sweden in 2005. The specific interest of the survey was a study of the trade-offs between salary and commuting time. For more detailed information on the data the reader is directed to Swärdh and Algers (2009).

As with any stated choice survey, the reliability of the data depends on respondents' limited ability to treat the attributes in isolation, i.e. there is a possibility that the sensitivity to salary changes will be to some extent influenced by the perceived effect that increases in travel time will have on increased travel costs. These issues, while important, are beyond the scope of the present paper, although we recognise the advantages of an approach jointly using stated preference and revealed preference data, such as in Dosman and Adamowicz (2006)[3]. The suitability of our data for the type of model discussed in this paper, despite not being traditional joint decision making data, stems from the fact that each person provides choices both for scenarios affecting only them and scenarios affecting both them and their partner. In fact, the absence of a negotiating process in such data, which would ideally require approaches such as discussed for example by Hensher et al. (2008), arguably avoids some of the issues arising in the application of such models to traditional joint choice data.

The study was conducted in two parts. First, each member of the household was asked to consider a choice between their current commute and one which would give them increased salaray in return for increased travel time. The survey thus looks at the willingness to accept (WTA) increased journey time in return for increased salary[4]. An example choice task for this first game is shown in Figure 1, where travel time is in minutes, and salary is in Swedish Kronor[5].

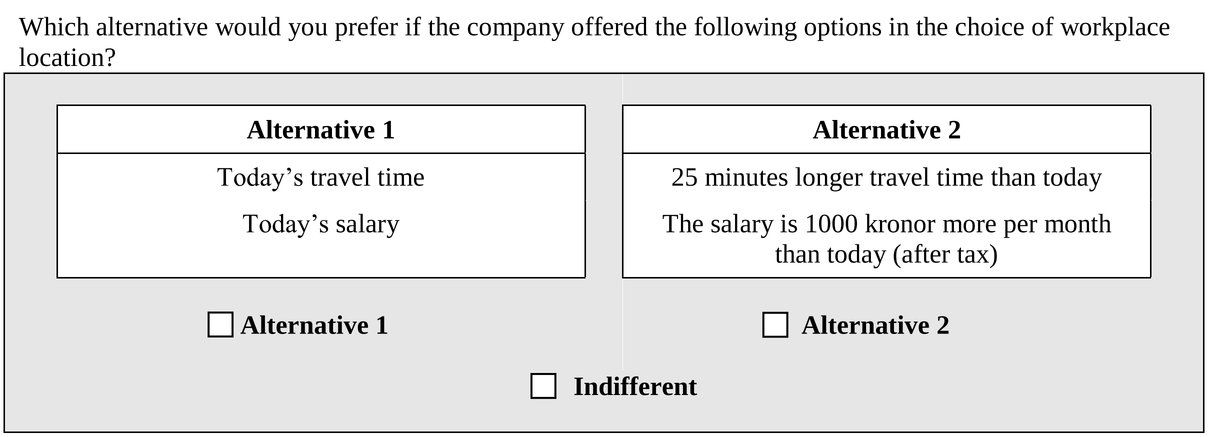

Fig. 1: Example of a stated choice scenario for game 1

Once the respondent had completed a series of these choice tasks they were then asked to complete

[3] We are grateful to an anonymous referee for highlighting this.

[4] The survey thus works with travel time per trip and salary per month. We acknowledge the different units of these two components and the potential shortcomings of this from a microeonomic theory perspective. However, from a behavioural perspective, salary is paid per month and travel time is experienced per journey, and this was the approach taken in the study - see also Swärdh and Algers (2009)

[5] The 2005 exchange is approximately £0.07 per SEK1.

the second part of the survey. In the second game, each respondent was asked in addition to consider the trade-off between increasing the length of time that it would take their partner to travel to work and an increase in their partner's monthly salary. An example choice task for this second game is shown in Figure 2. Crucially, the adjustments presented in this second task were not necessarily identical in proportion for the respondent and their partner.

Which alternative would you prefer if the company offered the following options in the choice of workplace location?

Alternative 1		**Alternative 2**	
You	Your partner	You	Your partner
Today's location *(Travel time and salary as today)*	Today's location *(Travel time and salary as today)*	25 minutes longer travel time than today The salary is 1000 kronor more per month than today (after tax)	10 minutes longer travel time than today The salary is 500 kronor more per month than today (after tax)

☐ **Alternative 1** ☐ **Alternative 2**

☐ **Indifferent**

Fig. 2: Example of a stated choice scenario for game 2

As can be seen from Figure 1 and Figure 2, each choice task contained two alternatives but the respondent was also given the opportunity to indicate indifference between the two options. For the purposes of the choice modelling analysis, this was coded as a third alternative. Each respondent was given four scenarios to complete in the first game, and an additional four or five tasks in the second game, depending on which version of the design was used. Within each household, the man and the woman by design usually received different versions of the survey. In total, responses were collected from 2,358 respondents, i.e. 1,179 couples. This provided us with a total of 20,041 observations.

3.2 Models not allowing for heterogeneity

3.2.1 Model specification

A number of different models were estimated, each time pooling the data from the choice tasks concerning only the household member completing the survey with the data from the choice tasks concerning both members. All models were estimated in Biogeme (Bierlaire, 2003). To recognise the repeated choice nature of the data, the standard errors in all models were computed using the panel specification of the sandwich matrix (cf. Daly and Hess, 2011).

For the first game, as shown in Figure 1, the observable component of the utility function for the three alternatives and individual n in choice scenario t is given by:

$$
\begin{aligned}
V_{nt1} &= \alpha_{1,1} + \beta_{\text{TT}}\text{TT}_{nt1} + \beta_{\text{L-Sal}}\text{L-Sal}_{nt1} \\
V_{nt2} &= \beta_{\text{TT}}\text{TT}_{nt2} + \beta_{\text{L-Sal}}\text{L-Sal}_{nt2} \\
V_{nt3} &= \alpha_{1,3}
\end{aligned} \tag{4}
$$

where β_{TT} and $\beta_{\text{L-Sal}}$ give the marginal utility coefficients for travel time (TT) and the logarithm of salary (L-Sal) - such a non-linear specification for salary produced superior results. Furthermore, $\alpha_{1,j}$ is the constant for alternative j in game 1, where, for identification reasons, we set $\alpha_{1,2} = 0$, thus estimating constants for the status quo alternative (alternative 1 above) and the "indifferent" *alternative* (alternative 3 above). We acknowledge that the treatment of the indifference alternative using a constant is simplistic in a random utility context, but a more detailed treatment was outside the scope of this analysis. For the travel time and salary attributes, the actual values were used, rather than the changes as presented in the survey, as this gave better model fit in the context of the non-linear specification for salary. When working with changes rather than absolute values, the solution would have been to interact the changes with the base level non-linearly[6].

For the second set of choices, as shown in Figure 2, (i.e., the *'joint'* game), the alternatives are now described by the travel time and salary for both partners, and the utilities are given by:

[6] We thank an anonymous referee for this comment.

$$
\begin{aligned}
V_{nt1} &= \nu \left[\alpha_{2,1} + \lambda \left(\beta_{\mathrm{TT}} \mathrm{TT}_{nt1} + \beta_{\text{L-Sal}} \text{L-Sal}_{nt1}\right)\right. \\
&\qquad\qquad \left. + \left(1-\lambda\right)\left(\beta_{\mathrm{TT}} \mathrm{TT}_{pt1} + \beta_{\text{L-Sal}} \text{L-Sal}_{pt1}\right)\right] \\
V_{nt2} &= \nu \left[\lambda \left(\beta_{\mathrm{TT}} \mathrm{TT}_{nt2} + \beta_{\text{L-Sal}} \text{L-Sal}_{nt2}\right)\right. \\
&\qquad\qquad \left. + \left(1-\lambda\right)\left(\beta_{\mathrm{TT}} \mathrm{TT}_{pt2} + \beta_{\text{L-Sal}} \text{L-Sal}_{pt2}\right)\right] \\
V_{nt3} &= \nu \alpha_{2,3}
\end{aligned}
\tag{5}
$$

This incorporates first a multiplication of the utility by ν, which gives the scale parameter for the second set of choices, with the scale for game 1 being normalised to 1. As in game 1, we estimate constants specific to game 2, namely $\alpha_{2,j}$, where $\alpha_{2,2} = 0$. The marginal utility coefficients are identical to those defined for Equation 4, while the associated attributes are now distinct for person n and their partner, indexed by p. The additional parameter λ refers to the weight that respondent n assigns to the circumstances affecting himself or herself, relative to those affecting their partner.

Whilst the specification in Equation 5 allows for respondent n to assign different weights to his/her own overall circumstances than those of his/her partner, it is conceivable that such differences also arise at the level of individual attributes, i.e. allowing for a greater disparity between the self and partner valuations for one attribute than for another. For this purpose, Equation 5 can be adapted to:

$$
\begin{aligned}
V_{nt1} &= \nu \left[\alpha_{2,1} + \lambda_{\mathrm{TT}} \beta_{\mathrm{TT}} \mathrm{TT}_{nt1} + \left(1-\lambda_{\mathrm{TT}}\right) \beta_{\mathrm{TT}} \mathrm{TT}_{pt1}\right. \\
&\qquad\qquad \left. + \lambda_{\text{L-Sal}} \beta_{\text{L-Sal}} \text{L-Sal}_{nt1} + \left(1-\lambda_{\text{L-Sal}}\right) \beta_{\text{L-Sal}} \text{L-Sal}_{pt1}\right] \\
V_{nt2} &= \nu \left[\lambda_{\mathrm{TT}} \beta_{\mathrm{TT}} \mathrm{TT}_{nt2} + \left(1-\lambda_{\mathrm{TT}}\right) \beta_{\mathrm{TT}} \mathrm{TT}_{pt2}\right. \\
&\qquad\qquad \left. + \lambda_{\text{L-Sal}} \beta_{\text{L-Sal}} \text{L-Sal}_{nt2} + \left(1-\lambda_{\text{L-Sal}}\right) \beta_{\text{L-Sal}} \text{L-Sal}_{pt2}\right] \\
V_{nt3} &= \nu \alpha_{2,3}
\end{aligned}
\tag{6}
$$

From Equation 6, it becomes clear that a corresponding specification could have been obtained without the λ parameters by instead using separate marginal utility coefficients for respondent n

and their partner p, as already alluded to in Section 2. We chose the above specification partly as it will facilitate interpretation in the models incorporating random heterogeneity, and avoids the need to specify correlation between β_n and β_p. The λ parameters now have even more importance than in Equation 5. Two views arise. They could be interpreted as differences the respondent perceives between his/her valuations of the attributes and those of his/her partner. Arguably more realistically, they could also be interpreted as the importance rating the respondent places on his/her own circumstances compared to those of their partner.

The specifications in Equations 4, 5 and 6 serve as the basis for the first three of our models. In particular:

Model 1 uses Equation 4 for the game 1 choices and Equation 5 for the game 2 choices, keeping λ fixed at 0.5, i.e. assuming that the decision maker gives equal weight to his/her partner.

Model 2 expands on model 1 by estimating λ.

Model 3 replaces Equation 5 with Equation 6, thus estimating separate λ parameters for travel time and salary.

3.2.2 Model results

The estimation results for the first three models are summarised in Table 1, where these models do not accommodate any heterogeneity across respondents, either deterministically or randomly. Looking at model 1, we see that all else being equal, there is some evidence of a preference for the status quo option (estimates for $\alpha_{1,1}$ and $\alpha_{2,1}$). The rate for the indifference alternative is below five percent, where we once again acknowledge the imperfect treatment of this alternative. The impact of increases in travel time is negative while the impact of increases in salary is positive, with the log-transform ensuring decreasing marginal returns. This model imposes the assumption that a respondent gives equal weight to both members of the household ($\lambda = 0.5$), while the scale parameter for the second game is not significantly different from the base of 1, suggesting no significant differences in the relative weight of the modelled and random utilities in the two games.

Looking next at model 2, which freely estimates λ, we note only a minor and not statistically significant improvement in model fit. This is in line with the estimate for λ changing only from

Tab. 1: Results: models 1 - 3

	Model 1 Equal weights		Model 2 Generic λ		Model 3 Attribute-specific λ	
	est.	t-rat.	est.	t-rat.	est.	t-rat.
$\alpha_{1,1}$	0.5370	9.67	0.5370	9.68	0.5370	9.67
$\alpha_{1,3}$	4.2100	2.76	4.2000	2.76	4.2100	2.76
$\alpha_{2,1}$	0.9210	7.03	0.9220	7.03	0.9240	7.04
$\alpha_{2,3}$	4.4000	2.78	4.3900	2.77	4.4000	2.77
β_{TT}	-0.0323	12.34	-0.0323	12.36	-0.0323	12.34
$\beta_{\text{L-Sal}}$	0.7330	4.91	0.7320	4.90	0.7330	4.91
λ	0.5	-	0.4870	12.98 *(0.35)*§	-	-
λ_{TT}	-	-	-	-	0.4730	11.89 *(0.54)*§
$\lambda_{\text{L-Sal}}$	-	-	-	-	0.5690	4.48 *(0.68)*§
ν	0.9240	11.42 *(0.94)*†	0.9240	11.42 *(0.94)*†	0.9230	11.43 *(0.95)*†
$\mathcal{L}\left(\hat{\beta}\right)$	-14,136.007		-14,135.945		-14,135.505	
$\bar{\rho}^2$	0.358		0.358		0.358	

† Note: t-rat. are relative to 1.
§ Note: t-rat. are relative to 0.5.

0.5 to 0.4870, where this change is not significant at the usual confidence levels. The remaining estimates remain unaffected.

A similar observation can be made for model 3, where the gains in fit obtained by allowing for attribute specific λ parameters are once again not significant at usual levels, and where neither weight parameter is significantly different from the base value of 0.5.

3.3 Models allowing for heterogeneity

3.3.1 Model specification

The three base models from Section 3.2 make the assumption of complete homogeneity across all respondents in all households for both the β and λ parameters. This assumption is gradually relaxed in the subsequent four models, which accommodate heterogeneity across respondents.

Model 4 expands on model 3 by accounting for deterministic heterogeneity by estimating separate β coefficients and separate λ coefficients for male and female respondents. This allows us to investigate whether there are any distinct differences by gender regarding how the members of the household dyad valued an increase in their own salary compared with how they valued an increase in their partner's salary, and in their willingness to accept a longer commute in return. This still equates to using Equation 4 and Equation 6, but with two sets of β and λ coefficients, relating to male and female respondents. It is important to note that this does not equate to using separate coefficients for the respondent and his/her partner in Equation 6.

In the final three models, we move to a specification accommodating random heterogeneity across respondents using Mixed Logit structures (see e.g. Train, 2009). Specifically, we still use separate parameters for male and female respondents, but now allow for additional random variation.

Model 5 expands on model 4 by allowing for additional random heterogeneity in the β parameters, using Lognormal distributions in a mixed logit model, where we allow for correlation between the travel time and salary coefficients, while still using separate coefficients for male and female respondents. In detail, and using the example of a female respondent, this equates to having:

$$\left\langle \ln\left(\beta_{f,\text{L-Sal}}\right), \ln\left(-\beta_{f,\text{TT}}\right)\right\rangle \sim MVN\left(\mu_{\beta_f}, \Omega_{\beta_f}\right), \tag{7}$$

such that the logarithms of the coefficients (with a sign change for the travel time coefficient) follow a multivariate Normal distribution, with mean $\mu_{\beta_f} = \left\langle \mu_{\ln\left(\beta_{f,\text{L-Sal}}\right)}, \mu_{\ln\left(-\beta_{f,\text{TT}}\right)}\right\rangle$, and covariance matrix $\Omega_{\beta_f} = \left\langle \sigma^2_{\ln\left(\beta_{f,\text{L-Sal}}\right)}, \sigma^2_{\ln\left(-\beta_{f,\text{TT}}\right)}, \sigma_{\ln\left(\beta_{f,\text{L-Sal}}\right),\ln\left(-\beta_{f,\text{TT}}\right)}\right\rangle$, where the first two terms relate to variances, and the third term is the covariance. In model estimation, this is achieved by using a Cholesky decomposition, which we return to below. A corresponding notation applies for male respondents. The distribution of random terms was carried out across households, where the panel specification ensured constant sensitivities for both individuals within a household across their choices (while still allowing for separate sensitivities for each of the individuals). For these models, the log-likelihood was simulated using 500 Halton draws (Halton, 1960).

Model 6 is a different generalisation of model 4 in that it allows for random heterogeneity in the λ

parameters, using Uniform distributions, with e.g.

$$\lambda_{f,\text{L-Sal}} \sim U\left[\lambda_{f,\mu_{\text{L-Sal}}} - \lambda_{f,s_{\text{L-Sal}}}, \lambda_{f,\mu_{\text{L-Sal}}} + \lambda_{f,s_{\text{L-Sal}}}\right], \tag{8}$$

so that $\lambda_{f,\text{L-Sal}}$ is uniformly distributed between $\lambda_{f,\mu_{\text{L-Sal}}} - \lambda_{f,s_{\text{L-Sal}}}$ and $\lambda_{f,\mu_{\text{L-Sal}}} + \lambda_{f,s_{\text{L-Sal}}}$.

Model 7 combines models 5 and 6, allowing for heterogeneity in both the β and λ parameters, using the same distributional assumptions as in these models, while still using separate parameters for male and female respondents.

3.3.2 Model results

We now turn our attention to models accommodating differences across respondents, where results for models 4 to 7 are summarised in Table 2. Model 4 expands on model 3 by allowing for differences between male and female respondents in the β and λ parameters, using subscripts m and f. This leads to an improvement in model fit by 4.11 units over model 3, which, at the cost of 4 additional parameters, is only significant at the 92% level. A detailed study of the results, using an asymptotic t-ratio for differences in parameters, reveals that the main differences arise in the β and λ parameters for travel time, although these differences are only significant at the 82% level for λ_{TT} and the 90% level for β_{TT}. Overall, this model would suggest only small differences between male and female respondents when accommodating deterministic heterogeneity alone.

The next step was to allow for random heterogeneity across respondents in the β parameters, where this is accommodated in model 5. As discussed before, we use multivariate Lognormal distributions, where $\mu_{\ln\left(\beta_{f,\text{L-Sal}}\right)}$ and $\mu_{\ln\left(-\beta_{f,\text{TT}}\right)}$ give the means of the underlying Normal distributions in the case of female respondents (where a corresponding notation with m applies to male respondents). We allow for correlation between the travel time and salary sensitivities and thus estimate three parameters for the Cholesky matrix, listed in the table as s terms. Hence, $|s_{11,\ln\left(\beta_{f,\text{L-Sal}}\right)}|$ gives the standard deviation for the underlying Normal distribution for $\ln\left(\beta_{f,\text{L-Sal}}\right)$, i.e. $\sigma_{\ln\left(\beta_{f,\text{L-Sal}}\right)}$ while the corresponding standard deviation for $\ln\left(-\beta_{f,\text{TT}}\right)$, i.e. $\sigma_{\ln\left(-\beta_{f,\text{TT}}\right)}$ is given by $\sqrt{s^2_{21,\ln\left(-\beta_{f,\text{TT}}\right)} + s^2_{22,\ln\left(-\beta_{f,\text{TT}}\right)}}$, with the covariance $\left(\sigma_{\ln\left(\beta_{f,\text{L-Sal}}\right),\ln\left(-\beta_{f,\text{TT}}\right)}\right)$ being equal to

Tab. 2: Results: models 4 - 7

	Model 4 Det. heterogeneity		Model 5 Random β		Model 6 Random λ		Model 7 Random β and λ	
	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
$\alpha_{1,1}$	0.5330	9.54	-0.8470	7.29	0.5420	9.62	-1.0300	6.91
$\alpha_{1,3}$	4.3100	2.74	-1.7700	5.59	4.7400	2.75	-1.9100	5.22
$\alpha_{2,1}$	0.9390	7.06	-1.0200	11.16	0.3050	1.92	-1.1200	10.32
$\alpha_{2,3}$	4.4800	2.77	-1.6300	5.84	5.4600	3.01	-1.7400	4.87
$\lambda_{f,\mu_{\text{L-Sal}}}$	0.5890	2.27 *(0.34)*[§]	0.5330	18.29 *(1.13)*[§]	-0.4870	1.56 *(3.16)*[§]	0.5340	12.67 *(0.81)*[§]
$\lambda_{f,s_{\text{L-Sal}}}$	-	-	-	-	4.0200	4.25	0.1230	0.89
$\lambda_{f,\mu_{\text{TT}}}$	0.5480	9.60 *(0.84)*[§]	0.6050	40.80 *(7.09)*[§]	0.5220	13.93 *(0.59)*[§]	0.6180	36.01 *(6.86)*[§]
$\lambda_{f,s_{\text{TT}}}$	-	-	-	-	0.0173	0.11	0.2620	5.66
$\lambda_{m,\mu_{\text{L-Sal}}}$	0.5720	2.77 *(0.35)*[§]	0.5580	16.76 *(1.74)*[§]	1.3700	3.89 *(2.47)*[§]	0.6260	13.00 *(2.61)*[§]
$\lambda_{m,s_{\text{L-Sal}}}$	-	-	-	-	3.4900	2.60	0.1720	3.68
$\lambda_{m,\mu_{\text{TT}}}$	0.4080	5.67 *(1.28)*[§]	0.5400	34.81 *(2.58)*[§]	0.4070	8.08 *(1.85)*[§]	0.5440	9.59 *(0.78)*[§]
$\lambda_{m,s_{\text{TT}}}$	-	-	-	-	0.8610	4.60	0.1250	0.26
$\beta_{f,\text{L-Sal}}$	0.7360	4.68	-	-	0.7770	4.56	-	-
$\beta_{f,\text{TT}}$	-0.0306	11.26	-	-	-0.0305	11.48	-	-
$\mu_{\ln(\beta_{f,\text{L-Sal}})}$	-	-	1.8300	22.42	-	-	1.4300	20.54
$\mu_{\ln(-\beta_{f,\text{TT}})}$	-	-	-1.6300	28.20	-	-	-1.6300	22.25
$s_{11,\ln(\beta_{f,\text{L-Sal}})}$	-	-	2.6600	39.22	-	-	2.6100	51.39
$s_{21,\ln(-\beta_{f,\text{TT}})}$	-	-	0.5860	11.82	-	-	0.5630	12.66
$s_{22,\ln(-\beta_{f,\text{TT}})}$	-	-	-0.3370	14.83	-	-	0.3060	16.15
$\beta_{m,\text{L-Sal}}$	0.7520	4.92	-	-	0.7900	4.74	-	-
$\beta_{m,\text{TT}}$	-0.0344	11.15	-	-	-0.0337	11.06	-	-
$\mu_{\ln(\beta_{m,\text{L-Sal}})}$	-	-	1.5500	13.26	-	-	1.6300	19.53
$\mu_{\ln(-\beta_{m,\text{TT}})}$	-	-	-1.6400	24.19	-	-	-1.5600	22.68
$s_{11,\ln(\beta_{m,\text{L-Sal}})}$	-	-	-2.7000	26.99	-	-	-2.6000	23.88
$s_{21,\ln(-\beta_{m,\text{TT}})}$	-	-	-0.6400	7.24	-	-	-0.5740	10.38
$s_{22,\ln(-\beta_{m,\text{TT}})}$	-	-	-0.3580	13.84	-	-	-0.4410	6.63
ν	0.9110	11.35 *(1.11)*[†]	2.0800	12.53 *(6.51)*[†]	1.6800	4.81 *(1.94)*[†]	1.9800	11.84 *(5.87)*[†]
$\mathcal{L}\left(\hat{\beta}\right)$	-14,131.392		-11,138.076		-14,007.193		-11,118.674	
$\bar{\rho}^2$	0.358		0.493		0.363		0.494	

[†] Note: t-rat. are relative to 1.
[§] Note: t-rat. are relative to 0.5.

$s_{11,\ln(\beta_{f,\text{L-Sal}})} s_{21,\ln(-\beta_{f,\text{TT}})}$. No sign constraint is imposed on any of the elements in the Cholesky matrix so as to allow for positive as well as negative covariances. The Cholesky parameters are obviously arbitrary depending on the order in which the coefficients are specified, whereas the required variance and covariance of the "true" parameters are unambiguous. For this reason, Table 3 also shows the implied distributions for the transformed parameters in the models 5 to 7.

Looking first at Table 2, we see that model 5 obtains a dramatic improvement in log-likelihood over model 4, with a hugely significant increase of 2,993.32 units at the cost of 6 additional parameters. This is a result of allowing for random heterogeneity as well as explicitly capturing the correlation across choices for the same respondent. The first observation to be made from the estimates for model 5 is that the constants for the first and third alternatives are now negative, possibly as a result of some of the behaviour previously captured by positive constants for the first and third alternative now being captured by the tails of the Lognormal distribution (remembering that the values for both the travel time and salary attributes are largest for the second alternative, which does not have a constant). We acknowledge that the tails of the lognormal distribution are long and entail high variances, but the distribution provided superior fit on this dataset and is in line with micro-economic theory when compared to unbounded alternatives. Additionally, the impact of the variances is reduced when looking at coefficient ratios in Section 3.4.

Turning to the λ parameters, we see that $\lambda_{f,\mu_{\text{TT}}}$ and $\lambda_{m,\mu_{\text{TT}}}$ are now significantly different from 0.5, while the differences between male and female respondents for $\lambda_{\mu_{\text{TT}}}$ are also statistically significant at high levels, with a t-ratio for differences of 3.13. Across all four λ parameters, we see an indication of greater weight being assigned to the respondent's attributes than to those of their partner.

All parameters relating to the lognormally distributed β coefficients are statistically significant. Using an asymptotic t-ratio for differences in parameters, we find that the differences between male and female respondents for the underlying mean for the salary distribution, $\mu_{\ln(\beta_{f,\text{L-Sal}})}$ and $\mu_{\ln(\beta_{m,\text{L-Sal}})}$, are significant with a confidence level of 97%. This observation, in line with a similar observation for the λ parameters above, suggests that the recovery of significant differences between male and female respondents is facilitated by additionally allowing for random heterogeneity across respondents. Finally, we see that the results for model 5 show significantly higher scale for game 2, i.e. the joint decisions, than for game 1. This was not the case in models 1 to 4, and could

Tab. 3: Analysis of random parameters for models 5 - 7

	Model 5 Random β		Model 6 Random λ		Model 7 Random β and λ	
	Male	Female	Male	Female	Male	Female
$\lambda_{\text{L-Sal}}$ (lower bound)	0.558	0.533	-2.120	-4.5070	0.454	0.411
$\lambda_{\text{L-Sal}}$ (upper bound)			4.860	3.5330	0.798	0.657
λ_{TT} (lower bound)	0.540	0.605	-0.454	0.5047	0.419	0.356
λ_{TT} (upper bound)			1.268	0.5393	0.669	0.880
$\mu_{\text{L-Sal}}$	180.37	214.39	0.79	0.78	149.90	125.97
$\sigma_{\text{L-Sal}}$	6,902.64	7,370.03	-	-	4,400.27	3,795.42
μ_{TT}	-0.25	-0.25	-0.03	-0.03	-0.27	-0.24
σ_{TT}	0.21	0.19	-	-	0.23	0.17
correlation ($\beta_{\text{L-Sal}}, \beta_{\text{TT}}$)	-0.17	-0.18	-	-	-0.18	-0.20

suggest that a failure to accommodate random variations in sensitivities led to an inability to adequately model the choices for game 2 in these earlier models, also reflected in our ability to now capture differences in the weights attached to a respondent and their partner. The finding of higher scale in more complex but still accessible choice tasks is not new (Caussade et al, 2005). A possible further interpretation for the higher scale in game 2 is that when being asked to make decisions on workplace location, a decision maker finds it easier to make an informed choice when having information on the effects for both household members. This would translate into more deterministic choices.

Looking at the implied heterogeneity patterns in Table 3, we observe very high levels of heterogeneity for the salary coefficients, with much more modest levels for the travel time coefficients[7]. There is negative correlation between the two coefficients, which is in line with expectations, where respondents who are more sensitive to salary are less sensitive to travel time, and vice versa. This is what drives the heterogeneity in the relative sensitivities between travel time and salary, where strong positive correlation would result in very low heterogeneity in the trade-offs. The actual implied differences in trade-offs between male and female respondents are studied in detail later.

Model 6 takes a different approach to model 5 by allowing for heterogeneity in the λ parameters

[7] While $\mu_{\ln\left(\beta_{f,\text{L-Sal}}\right)}$ in Table 2 relates to the mean of the underlying Normal distribution for the salary coefficient for female respondents, $\mu_{\text{L-Sal}}$ represents the resulting mean of the Lognormal distribution, with $\sigma_{\text{L-Sal}}$ giving the resulting standard deviation. The means and standard deviations for the Lognormal distribution can be obtained as simple transforms of the parameters for the underlying Normal distribution reported in Table 2, using the formulae reported in Train (2009, page 150).

rather than the β parameters, where Uniform distributions are used, with e.g. $\lambda_{\mathrm{f,L\text{-}Sal}}$ having a mean of $\lambda_{f,\mu_{\text{L-Sal}}}$, with Uniform variation between $\lambda_{f,\mu_{\text{L-Sal}}} - \lambda_{f,s_{\text{L-Sal}}}$ and $\lambda_{f,\mu_{\text{L-Sal}}} + \lambda_{f,s_{\text{L-Sal}}}$. This model obtains an improvement in log-likelihood by 124.19 units over model 4, which is statistically significant at the cost of 4 additional parameters, but is clearly far more modest than the improvement obtained by model 5. As in model 5, we again see heightened scale for game 2. However, a further inspection of the estimates (see Table 3) shows that with the exception of $\lambda_{f,\mathrm{TT}}$, the range of the λ parameters falls outside the $[0, 1]$ boundary, where, for $\lambda_{f,\text{L-Sal}}$, we even obtain a negative mean. As noted earlier, a number of interpretations have been put forward for such estimates, but we believe that at least in some cases, this is a result of confounding with other heterogeneity, a point we investigate further in model 7. Additionally, in the present case, negative λ parameters would lead to a change in the sign of the marginal utility coefficients, which is clearly nonsensical. A further potential reason for sign violations of the range of weight parameters could be where the true distribution is asymmetrical while the analyst attempts to fit a symmetrical distribution. However, the results from model 7 seem to rather point in the direction of unaccounted for heterogeneity in the marginal utility coefficients.

Model 7 presents a generalisation of both model 5 and model 6. In comparison with model 5, we obtain gains in log-likelihood by 19.40 which is statistically significant, at the cost of 4 additional parameters. Similarly, model 7 obtains a hugely significant improvement in log-likelihood by 2,888.52 units over model 6, at the cost of 6 additional parameters. This shows the benefit of allowing jointly for heterogeneity in β and λ, although some of the gains over model 5 could be the result of the more flexible distributional assumptions for the marginal utility coefficients in game 2 (Uniform multiplying a Lognormal, instead of a Lognormal alone). We can see from Table 3 that jointly accommodating heterogeneity in β and λ leads to reductions in the levels of heterogeneity (e.g. the coefficient of variation for salary for male respondents drops from 38.27 to 29.35), albeit that the tails of the Lognormal clearly remain quite influential. As was the case in model 5, the constants for the first and third alternative are once again negative. The parameters for the lognormally distributed β coefficients again all attain high levels of significance, although it needs to be recognised that these relate to the parameters of the underlying Normal distribution and that the significance levels may be different for the transformed parameters (i.e. on the Lognormal scale). Crucially, in contrast with model 6, all λ parameters now have a range that is strictly within the $[0, 1]$ interval (cf. Table 3). This final model is also more successful in retrieving significant differences between male and female respondents, in line with similar observations for model 5 - for example, we find that the differences between male and female respondents for the underlying mean for the salary distribution, $\mu_{\ln\left(\beta_{f,\text{L-Sal}}\right)}$ and $\mu_{\ln\left(\beta_{m,\text{L-Sal}}\right)}$, are significant with a confidence level of 99%.

3.4 Implied trade-offs

As a next step in our comparison between the different models, we now look at relative valuations of the two attributes. The context of the survey was a study of the willingness by respondents to accept higher travel time in return for higher salary, and as such, the focus in this section is specifically on that ratio, as opposed to the willingness to accept lower salary in return for shorter travel times, which would be similar in meaning to the widely used value of travel time savings.

The calculation of the ratios between the two coefficients is complicated by the use of the log-transform for salary in all models, meaning that the WTA reduces with increasing income. This implies, quite logically, that, as the marginal benefit of increased salary is decreasing, i.e. at higher salaries, a respondent becomes less sensitive to salary increases, this yields a lower willingness to accept increased travel time in return for salary increases. In a model with fixed coefficients only, the trade-off would be given by $\frac{\beta_{\text{L-Sal}}}{\beta_{\text{TT}}} \cdot \frac{1}{\text{Sal}}$, i.e. the trade-off is divided by the salary and we get a lower willingness to accept travel time increases in return for salary reductions for respondents with higher salary[8]. By thinking about the inverse of this ratio, we can see that the relative importance of time against money increases as salary increases, which is consistent with the usual finding of a value of time increasing with salary.

Given the above non-linearities, our analysis calculated individual WTA values for each SP observation in the data, using the salary for the chosen alternative, and our results look at the distribution of the resulting WTA measures in the sample population. The decision to work the WTA out at the chosen salary rather than at the status quo or current salary is based on a desire to compute the WTA in the stated choice data rather than in the RP market. However, it should be

[8] Looking at model 1, we have that the ratio between the log-salary and time coefficients is equal to 22.69. This then needs to divided by a respondent's salary to get the implied WTA. For example, the lowest male salary is SEK3,750, giving a willingness to accept 0.006 minutes per additional Krona. For a respondent at the highest male salary, in this case SEK75,000, the WTA is much lower, at 0.0003 minutes per additional Krona.

Tab. 4: Results: trade-offs

WTA extra mins per trip for 1,000K extra a month

	Female respondents					
	Self			**Partner**		
	mean	s.d.	cv	mean	s.d.	cv
Model 1	1.1016	0.81	0.74	0.7927	0.49	0.62
Model 2	1.1001	0.81	0.74	0.7916	0.49	0.62
Model 3	1.3251	0.98	0.74	0.6483	0.40	0.62
Model 4	1.2549	0.93	0.74	0.7640	0.47	0.62
Model 5	12.3723	122.79	9.92	11.9482	150.41	12.59
Model 6		*undefined*			*undefined*	
Model 7	9.5305	105.23	11.04	7.6079	88.15	11.59
	Male respondents					
	Self			**Partner**		
	mean	s.d.	cv	mean	s.d.	cv
Model 1	0.7897	0.48	0.61	1.1200	0.86	0.76
Model 2	0.7887	0.48	0.61	1.1184	0.85	0.76
Model 3	0.9500	0.58	0.61	0.9160	0.70	0.76
Model 4	1.0666	0.65	0.61	0.7800	0.60	0.76
Model 5	7.7722	83.28	10.71	10.2491	109.88	10.72
Model 6		undefined			undefined	
Model 7	8.6593	88.47	10.22	8.4555	95.26	11.27

noted that this had a negligible effect on results. Overall WTA measures would have been higher by just 1.3% when using the status quo income (which is on average lower than the chosen income), with the standard deviation of the WTA measures increasing by 4.6% overall.

The calculation becomes somewhat more complicated once we introduce λ parameters as well as deterministic and random heterogeneity across respondents. Here, the mean and standard deviations are calculated analytically rather than using simulation, which would be unreliable due to the long tails of the Lognormal distribution. An important issue arises in model 6. The fact that the distribution of the λ parameters falls outside the $[0, 1]$ range means that the moments of the resulting WTA distribution are undefined (cf. Daly et al., 2012), and as such are not reported. This is a further reason for attempting to ensure constant signs across respondents in the λ parameters, a point seemingly not recognised in earlier work.

A number of key observations can be made from the results in Table 4. Accommodating random heterogeneity across respondents in the β parameters obviously leads to a very significant increase in heterogeneity in the WTA measures, whereas the heterogeneity in the initial models is merely a result of the non-linear specification (using the logarithm of salary). At the same time, we also see a significant increase in the mean WTA measures, leading to more realistic values than was the case in the first four models by bringing them closer to common value of time findings.

Focussing on the results from model 7, which gave the best overall performance, we can see that for female respondents, the WTA measures for the respondents themselves are higher than those they assign to their male partners. Although female respondents assign more weight to their partner's salary than his travel time, which would imply higher WTA, the actual salary for male respondents is higher in this sample, leading to lower WTA measures. Male respondents on the other hand assign more weight to their partner's travel time than to her salary, which would lead to low WTA measures, but this is compensated for by the lower salary for female respondents in the data, meaning that the final WTA measures assigned by male respondents to themselves and their partner are very similar.

4 Conclusions

This paper has focussed on the issue of the representation of heterogeneity in choice models that are either estimated on data from joint decisions or data on decisions made by a single person but affecting multiple individuals. Our empirical example has focussed on the latter.

A number of central ideas are put forward in the paper, and tested in an empirical study using a stated choice dataset in which each partner was asked to evaluate scenarios leading to changes in travel time and salary for both themselves and their partner.

Firstly, we argue that differences in weights assigned to individual partners of a household may vary across attributes. Our results show that the weights respondents assign to their partners do indeed vary across attributes, although such differences are only properly retrieved when allowing for heterogeneity in the marginal utility coefficients[9]. For example, using an asymptotic t-ratio for differences in parameters, we find significant differences between the mean female allocation of salary and travel time weights, $\lambda_{f,\mu_{\text{TT}}}$ and $\lambda_{f,\mu_{\text{L-Sal}}}$ respectively, in both model 5 and model 7, with a confidence level of 92% applying to the differences in model 7.

Secondly, we argue that there is scope for significant heterogeneity across respondents in under-

[9] Note that efforts to study differences between λ_{TT} and $\lambda_{\text{L-Sal}}$ were only moderately successful in models 3 and 4.

lying sensitivities as well as the relative weights assigned to themselves and their partners. This is once again confirmed in the empirical example, showing significant improvements in model fit when allowing for random heterogeneity in the β parameters, and to a lesser extent in the λ parameters. We also retrieve differences between male and female respondents in both sets of parameters, but here there is evidence that such differences can only be adequately captured if simultaneously accommodating random variations.

Thirdly, and most importantly, we argue that there is potentially significant scope for confounding between heterogeneity in marginal sensitivities and heterogeneity in bargaining or weight parameters. Additionally, there is a risk of inappropriate assumptions for the distribution of randomly distributed bargaining or weight parameters leading to misguided results and interpretations. These claims are strongly supported by the evidence from model 6. This model shows that only allowing for heterogeneity in λ without accounting for heterogeneity in β leads to overstated heterogeneity in the former, along with suggesting a significant share of the distribution for λ falling outside the conventional $[0, 1]$ range. While arguments have been put forward to justify such values, we argue here that an incomplete or inappropriate treatment of heterogeneity in the β parameters may exacerbate such problems; a claim entirely supported by the differences in results between model 6 and model 7, notwithstanding the slightly different role for λ in our models. It may also play a role in results showing a dominant role for one partner, e.g. as in Dosman and Adamowicz (2006). Clearly, it is also crucial not to use distributional assumptions that would a priori postulate the presence of such values, such as in the use of a normally distributed λ parameter (cf. Beharry-Borg et al., 2009); here the same argument applies as for marginal utility coefficients with strong a priori sign expectations (cf. Hess et al., 2005). In a specification such as used here, a negative λ parameter would also lead to sign violations for the marginal utility coefficients.

The greater ability of retrieving heterogeneity in the λ parameters when additionally accommodating random heterogeneity in the marginal utility coefficients is also highlighted in Table 5, which again shows the problems arising with model 6 due to its failure to account for such heterogeneity in β while allowing for heterogeneity in λ.

In terms of actual empirical findings for the data at hand, there is evidence of significant heterogeneity across respondents in their own trade-offs between salary and travel time, as well as the weight they assign for those two attributes for their partner. Most of this heterogeneity is

Tab. 5: Results: weight parameters

	Travel time					
	Female			Male		
	Lower bound	Mean	Upper bound	Lower bound	Mean	Upper bound
Model 1	-	0.5	-	-	0.5	-
Model 2	-	0.4870	-	-	0.4870	-
Model 3	-	0.4730	-	-	0.4730	-
Model 4	-	0.5480	-	-	0.4080	-
Model 5	-	0.6050	-	-	0.5400	-
Model 6	0.4507	0.5220	0.5933	-0.4540	0.4070	1.2680
Model 7	0.3560	0.6180	0.8800	0.4190	0.5440	0.6690
	Salary					
	Female			**Male**		
	Lower bound	Mean	Upper bound	Lower bound	Mean	Upper bound
Model 1	-	0.5	-	-	0.5	-
Model 2	-	0.4870	-	-	0.4870	-
Model 3	-	0.5690	-	-	0.5690	-
Model 4	-	0.5890	-	-	0.5720	-
Model 5	-	0.5330	-	-	0.5580	-
Model 6	-4.5070	-0.4870	3.5330	-2.1200	1.3700	4.8600
Model 7	0.4110	0.5340	0.6570	0.4540	0.6260	0.7980

random, but some is also linked to differences between men and women. Here, there is evidence that male respondents give more weight to their partner's travel time than to her salary, with the opposite applying to female respondents. These differences do not translate directly into the WTA patterns though, given the non-linear valuation of increases in salary and the higher overall salary for male respondents.

There is significant scope for future work. This includes attempts to validate our findings on other data, looking into the impact of heterogeneity assumptions in a more traditional joint decision making context, as well as studies across a range of topic areas, including leisure and non-leisure activities. Future work should also concentrate more on linking heterogeneity in λ to underlying respondent characteristics, where the main emphasis thus far has been on income, but where scope also exists to study the impact of gender roles, the relative levels of education of each of the household members, and their employment status and patterns. In general, greater effort should go into explaining heterogeneity in both λ and β in such a deterministic manner, but in the present case, gender was the main discriminator. Similarly, there is scope for testing non-linear formulations for the weight parameters in future work, where in the present paper, we restricted ourselves to a standard linear specification.

Acknowledgements

The authors are indebted to Danny Campbell for important contributions in the early stages of the work leading to this paper, and for his guidance to the first author during her PhD studies. The authors are also grateful to Staffan Algers for making the survey data available, and to Amanda Stathopoulos for translating the survey from Swedish - without their help, this work would not have been possible. The authors would also like to acknowledge the input of Nesha Beharry-Borg in early stages of this research. Earlier versions of this paper were presented at the European Transport Conference 2010 held in Glasgow and the Transportation Research Board (TRB) 91st Annual Meeting 2012 held in Washington, D.C. The authors are grateful for the comments received at these conferences, which provided insightful suggestions for revisions, as well as the comments of three anonymous referees received on an earlier version.

References

Adamowicz, W., Hanemann, M., Swait, J., Johnson, R., Layton, D., Regenwetter, M., Reimer, T., Sorkin, R., 2005. Decision strategy and structure in households: a “groups” perspective. Marketing Letters 16 (3/4), 387–399.

Aribarg, A., Arora, N., Bodur, H. O., 2002. Understanding the role of preference revision and concession in group decisions. Journal of Marketing Research 39 (3), 336–349.

Arora, N., Allenby, G. M., 1999. Measuring the influence of individual preference structures in group decision making. Journal of Marketing Research 36 (4), 476–487.

Bateman, I., Munro, A., 2005. An experiment on risky choice amongst households. The Economic Journal 115 (502), C176–C189.

Beharry-Borg, N., Hensher, D. A., Scarpa, R., 2009. An analytical framework for joint vs separate decisions by couples in choice experiments: the case of coastal water quality in Tobago. Environmental and Resource Economics 43 (1), 95–117.

Bierlaire, M., 2003. BIOGEME: a free package for the estimation of discrete choice models. Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.

Browning, M., Chiappori, P. A., 1998. Efficient intra-household allocations: a general characterization and empirical tests. Econometrica 66 (6), 1241–1278.

Caussade, S., Ortúzar, J. de D., Rizzi, L., Hensher, D.A., 2005. Assessing the influence of design dimensions on stated choice experiment estimates. Transportation Research Part B 39(7), 621–640

Daly, A. J., Hess, S., 2011. Simple methods for panel data analysis. Paper presented at the 90^{th} Annual Meeting of the Transportation Research Board, Washington, D.C.

Daly, A., Hess, S., Train, K., 2012. Assuring finite moments for willingness to pay estimates from random coefficients models. Transportation 39 (1), 19–31.

Dellaert, B. G. C., Prodigalidad, M., Louviere, J. J., 1998. Family members' projections of each other's preference and influence: a two-stage conjoint approach. Marketing Letters 9 (2), 135–145.

Dosman, D., Adamowicz, W., 2006. Combining stated and revealed preference data to construct an empirical examination of intrahousehold bargaining. Review of Economics of the Household 4 (1), 15–34.

Halton, J., 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numerische Mathematik 2, 84–90.

Hensher, D. A., Rose, J. M., Black, I., 2008. Interactive agency choice in automobile purchase decisions: the role of negotiation in determining equilibrium choice outcomes. Journal of Transport Economics and Policy 42 (2), 269–296.

Hess, S., Bierlaire, M., Polak, J. W., 2005. Estimation of value of travel-time savings using mixed logit models. Transportation Research Part A: Policy and Practice 39 (2/3), 221–236.

Katz, E., 1997. The intra-household economics of voice and exit. Feminist Economics 3 (3), 25–46.

Lampietti, J., 1999. Do husbands and wives make the same choices? Evidence from Northern Ethiopia. Economics Letters 62 (2), 253–260.

Marcucci, E., Stathopoulos, A., Rotaris, L., Danielis, R., 2011. Comparing single and joint preferences: a choice experiment on residential location in three member households. Environment and Planning A 43 (5), 1209–1225.

Munro, A., 2009. Introduction to the special issue: things we do and don't understand about the household and the environment. Environmental and Resource Economics 43 (1), 1–10.

Myers, D. G., Lamm, H., 1976. The group polarization phenomenon. Psychological Bulletin 83 (4), 602–627.

Rao, V. R., Steckel, J. H., 1991. A polarization model for describing group preferences. The Journal of Consumer Research 18 (1), 108–118.

Steckel, J. H., Corfman, K. P., Curry, D. J., Gupta, S., Shanteau, J., 1991. Prospects and problems in modeling group decisions. Marketing Letters 2 (3), 231–240.

Swärdh, J., Algers, S., 2009. Willingness to accept commuting time for yourself and for your spouse: empirical evidence from Swedish stated preference data. Working Papers 2009:5, Swedish National Road & Transport Research Institute (VTI).

Train, K., 2009. Discrete choice methods with simulation, Second Edition. Cambridge University Press.

Vermeulen, F., 2002. Collective household models: principles and main results. Journal of Economic Surveys 16 (4), 533–564.